

**PERBANDINGAN ANALISIS KLASIFIKASI NASABAH KREDIT
MENGUNAKAN REGRESI LOGISTIK BINER DAN CART
(*CLASSIFICATION AND REGRESSION TREES*)**



SKRIPSI

Oleh :

AGUNG WALUYO

24010210141020

**JURUSAN STATISTIKA
FAKULTAS SAINS DAN MATEMATIKA
UNIVERSITAS DIPONEGORO
SEMARANG**

2015

**PERBANDINGAN ANALISIS KLASIFIKASI NASABAH KREDIT
MENGUNAKAN REGRESI LOGISTIK BINER DAN CART
(*CLASSIFICATION AND REGRESSION TREES*)**

**AGUNG WALUYO
24010210141020**

Skripsi

Diajukan Sebagai Syarat untuk Memperoleh Gelar Sarjana Sains
pada Jurusan Statistika

**JURUSAN STATISTIKA
FAKULTAS SAINS DAN MATEMATIKA
UNIVERSITAS DIPONEGORO
SEMARANG
2015**

HALAMAN PENGESAHAN I

Judul : Perbandingan Analisis Klasifikasi Nasabah Kredit Menggunakan
Regresi Logistik Dan CART (*Classification And Regression
Trees*)

Nama : Agung Waluyo

NIM : 24010210141020

Jurusan : Statistika

Telah diujikan pada sidang Tugas Akhir tanggal 18 Februari 2015 dan dinyatakan
lulus pada tanggal 24 Februari 2015.

Semarang, Februari 2015

Mengetahui,

Ketua Jurusan Statistika

FSM UNDIP,



Dra. Dwi Ispriyanti, M.Si
NIP. 195709141986032001

Ketua Panitia Penguji

Ujian Tugas Akhir,



Drs. Sudarno, M.Si

NIP. 196407091992011001

HALAMAN PENGESAHAN II

Judul : Perbandingan Analisis Klasifikasi Nasabah Kredit Menggunakan
Regresi Logistik Dan CART (*Classification And Regression
Trees*)

Nama : Agung Waluyo

NIM : 24010210141020

Jurusan : Statistika

Telah diujikan pada sidang Tugas Akhir tanggal 18 Februari 2015 dan dinyatakan
lulus pada tanggal 24 Februari 2015.

Semarang, Februari 2015

Pembimbing I



Moch. Abdul Mukid, S.Si, M.Si

NIP. 197808172005011001

Pembimbing II



Triastuti Wuryandari, S.Si, M.Si

NIP. 197109061998032001

KATA PENGANTAR

Puji syukur kehadiran Allah SWT yang telah memberikan rahmat, hidayah, dan karunia-Nya sehingga penulis dapat menyelesaikan penulisan skripsi dengan judul “Perbandingan Analisis Klasifikasi Nasabah Kredit Menggunakan Regresi Logistik Biner Dan CART (*Classification And Regression Trees*)”.

Penulis menyadari skripsi ini tidak akan dapat diselesaikan tanpa bantuan dari berbagai pihak. Oleh karena itu, penulis ingin menyampaikan terima kasih kepada:

1. Ibu Dra. Dwi Ispriyanti, M.Si selaku Ketua Jurusan Statistika Fakultas Sains dan Matematika Universitas Diponegoro.
2. Bapak Moch. Abdul Mukid, S.Si, M.Si dan Ibu Triastuti Wuryandari, S.Si, M.Si selaku dosen pembimbing I dan II yang telah berkenan meluangkan waktu dalam memberikan masukan, arahan, dan bimbingan kepada penulis.
3. Bapak dan Ibu Dosen Jurusan Statistika yang sudah memberikan ilmu selama ini
4. Semua pihak yang telah membantu dalam penulisan skripsi ini yang tidak dapat disebutkan satu per satu.

Penulis menyadari bahwa skripsi ini masih jauh dari kesempurnaan. Oleh karena itu kritik dan saran yang sifatnya membangun sangat penulis harapkan demi kesempurnaan penulisan selanjutnya.

Semarang, Februari 2015

Penulis

ABSTRAK

Kredit merupakan aset paling besar yang dikelola bank dan juga merupakan kontributor yang paling dominan terhadap pendapatan bank. Debitur dalam membayar kredit ke bank mungkin lancar atau macet. Ada hubungan variabel yang mempengaruhi seorang debitur lancar atau macet dalam membayar kredit. Penelitian ini bertujuan untuk mengidentifikasi variabel-variabel yang mempengaruhi status kredit seorang debitur. Variabel-variabel yang digunakan dalam penelitian ini adalah jenis kelamin, jenjang pendidikan, jenis pekerjaan, status pernikahan dan pendapatan. Metode analisis yang digunakan meliputi analisis Regresi Logistik Biner dan CART (*classification and regression trees*). Akurasi pengklasifikasian dari kedua metode tersebut akan dibandingkan. Berdasarkan hasil penelitian dengan metode regresi logistik biner menunjukkan bahwa variabel yang berpengaruh terhadap status kredit debitur adalah pendapatan dengan ketepatan klasifikasi 80%. Sedangkan Hasil CART (*classification and regression trees*) berupa pohon keputusan menunjukkan jenis pekerjaan terpilih sebagai pemecah simpul akar, dengan akurasi klasifikasi sebesar 81%.

Kata kunci : Status Kredit, Regresi Logistik Biner, CART

ABSTRACT

Credit is the greatest asset managed by the bank and also the most dominant contributor to the bank's revenue. Debtor to pay credit to the bank may smoothly or jammed. There is a relationship variables that affect a debtor smoothly or jammed in paying credit. This study aims to identify the variables that affect a debtor's credit status. The variables used in this study were gender, education level, occupation, marital status and income. Analytical methods used include Binary Logistic Regression analysis and CART (classification and regression trees). Classification accuracy of the two methods will be compared. Based on the research results of binary logistic regression showed that the variables that affect the debtor's credit status is revenue with 80% classification accuracy. While the results of CART (classification and regression trees) in the form of a decision tree shows the type of work chosen as the root node splitting, with a classification accuracy of 81%.

Keywords: credit status, logistic regression, CART

DAFTAR ISI

	Halaman
HALAMAN JUDUL	i
HALAMAN PENGESAHAN I	ii
HALAMAN PENGESAHAN II	iii
KATA PENGANTAR	iv
ABSTRAK	v
ABSTRACT	vi
DAFTAR ISI	vii
DAFTAR TABEL	x
DAFTAR GAMBAR	xi
DAFTAR LAMPIRAN	xii
BAB I PENDAHULUAN	
1.1 Latar Belakang	1
1.2 Permasalahan	3
1.3 Batasan Masalah	3
1.4 Tujuan Penulisan	3
BAB II TINJAUAN PUSTAKA	
2.1 Pengertian Bank	4
2.2 Pengertian Kredit	6
2.3 Pengertian Nasabah	8
2.4 Model Regresi Logistik Biner	9
2.4.1 Estimasi Parameter	10

2.4.2 Uji Signifikansi.....	13
2.4.3 Pengukuran Ketepatan Hasil Klasifikasi	15
2.5 CART (<i>Classification And Regresion Tress</i>)	16
2.6 Struktur atau Bentuk Pohon Klasifikasi CART.....	17
2.7 <i>Binary Recursive Partitioning</i>	19
2.8 Langkah-Langkah Kerja CART	20
2.8.1 Proses Pemecahan Simpul.....	21
2.8.2 Pelabelan kelas (<i>Class Assignment</i>).....	23
2.8.3 Proses Penghentian Pemecahan.....	24
2.8.4 Proses Pemangkasan Pohon	24
BAB III METODOLOGI PENELITIAN	
3.1 Data	28
3.2 Metode Analisis	29
3.3 Diagram Alir	31
BAB IV HASIL DAN PEMBAHASAN	
4.1 Statistika Deskriptif	32
4.2 Analisis Status Kredit Bank dengan Regresi Logistik Biner	38
4.2.1 Model Awal.....	38
4.2.2 Uji Rasio Likelihood	38
4.2.3 Uji Wald	39
4.2.4 Uji Rasio Likelihood Kedua.....	41
4.2.5 Uji Wald Kedua.....	41
4.2.6 Uji Kesesuaian Model	42
4.2.7 Model Akhir	43

4.2.8 Interpretasi Model Regresi Logistik Biner	43
4.2.9 Ketepatan Klasifikasi Menggunakan Regresi Logistik Biner	44
4.3 Analisis Status Kredit Bank dengan CART (<i>Classification And Regression Tress</i>)	46
4.3.1 Proses <i>Splitting</i> Simpul	47
4.3.2 Pelabelan Kelas (<i>Class Assignment</i>)	49
4.3.3 Proses Penghentian Pemecahan (<i>Stop The Split</i>)	50
4.3.4 Proses Pemangkasan Pohon (<i>Trees Prunning Process</i>).....	50
4.3.5 Interpretasi Pohon Klasifikasi.....	56
4.3.2 Ukuran Ketepatan Klasifikasi	59
BAB V KESIMPULAN	60
DAFTAR PUSTAKA	60

DAFTAR TABEL

	Halaman
Tabel 1 Matriks Konfusi	16
Tabel 2 Variabel Data Debitur	28
Tabel 3 Hasil Uji Wald Pertama	40
Tabel 4 Hasil Uji Wald Kedua	42
Tabel 5 Estimasi Nilai Peluang Model.....	44
Tabel 6 Nilai <i>Odds Ratio</i>	44
Tabel 7 Tabel Ketepatan Klasifikasi Regresi Logistik Biner.....	45
Tabel 8 Nilai <i>Goodness of Split</i> Calon Pemecah	48
Tabel 9 Tabel Ketepatan Klasifikasi CART	59

DAFTAR GAMBAR

	Halaman
Gambar 1 Pohon Klasifikasi Biner CART	17
Gambar 2 Proses Partisi	19
Gambar 3 Pohon Klasifikasi T	25
Gambar 4 Cabang T t_2	26
Gambar 5 Pohon Klasifikasi T- T t_2	26
Gambar 6 Diagram Status Kredit Bank	32
Gambar 7 Diagram Status Kredit Berdasarkan Jenis Kelamin	33
Gambar 8 Diagram Status Kredit Bank Berdasarkan Jenjang Pendidikan	34
Gambar 9 Diagram Status Kredit Bank Berdasarkan Status Pernikahan	35
Gambar 10 Diagram Status Kredit Bank Berdasarkan Jenis Pekerjaan	36
Gambar 11 Diagram Status Kredit Bank Berdasarkan Pendapatan	37
Gambar 12 Variabel Pemecah Terbaik	49
Gambar 13 Simpul akhir 49 dan 50	50
Gambar 14 Pohon Klasifikasi Sebelum Dipangkas	51
Gambar 15 Pohon Klasifikasi Setelah Dipangkas	56

DAFTAR LAMPIRAN

	Halaman
Lampiran 1 Data Status Kredit Nasabah Sebuah Bank.....	62
Lampiran 2 Tabulasi Silang Variabel Bebas terhadap Variabel Tak Bebas	63
Lampiran 3 Regresi Logistik Biner Kondisi Pertama	65
Lampiran 4 Regresi Logistik Biner Dengan Menggunakan Variabel yang Signifikan	70
Lampiran 5 CART (<i>Classification And Regression Trees</i>) Sebelum <i>Prunning</i>	72
Lampiran 6 CART (<i>Classification And Regression Trees</i>) Setelah <i>Prunning</i> ...	78
Lampiran 7 Tabel <i>Chi-square</i>	80

BAB I

PENDAHULUAN

1.1 Latar Belakang

Bank menurut Undang-undang RI nomor 10 tahun 1998 tanggal 10 November 1998 tentang perbankan adalah Badan usaha yang menghimpun dana dari masyarakat dalam bentuk simpanan dan menyalurkan kepada masyarakat dalam bentuk kredit dan atau bentuk-bentuk lainnya dalam rangka meningkatkan taraf hidup rakyat banyak. Bank merupakan perusahaan yang bergerak dalam bidang keuangan, artinya usaha perbankan selalu berkaitan dengan masalah di bidang keuangan. Kegiatan menghimpun dan menyalurkan dana merupakan kegiatan pokok perbankan, sedangkan kegiatan memberikan jasa-jasa bank hanyalah merupakan kegiatan pendukung. Fungsi utama bank dalam suatu perekonomian adalah untuk memobilisasi dana masyarakat, secara tepat dan cepat menyalurkan dana tersebut kepada pengguna atau investasi yang efektif dan efisien (Kasmir, 2002).

Penerapan prinsip kehati-hatian oleh bank diantaranya diimplementasikan melalui kemampuan bank untuk mengelola portofolio kredit yang dimiliki sehingga resiko yang berpotensi untuk terjadi (*credit risk*) dapat diukur dan dikontrol. Kredit merupakan asset yang paling besar yang dikelola bank dan juga merupakan kontributor yang paling dominan terhadap pendapatan bank. Model skor kredit merupakan alat bantu dalam melakukan analisa kelayakan kredit berguna sebagai langkah awal dalam mengurangi resiko terjadinya kegagalan pemenuhan kewajiban oleh debitur (Pandia, 2012).

Metode klasifikasi dapat dilakukan dengan pendekatan parametrik dan nonparametrik. Dalam pendekatan parametrik terdapat beberapa metode klasifikasi, salah satunya adalah analisis regresi logistik. Menurut Hosmer (1989), metode regresi logistik adalah suatu metode analisis statistika yang mendeskripsikan hubungan antara variabel tak bebas yang memiliki dua kategori atau lebih dengan satu atau lebih variabel bebas berskala kategori atau interval. Variabel kategorik yaitu variabel dengan skala ukur nominal atau ordinal. Regresi logistik biner adalah salah satu metode statistika yang sering digunakan untuk mengklasifikasikan sejumlah pengamatan dengan variabel tak bebas biner ke dalam beberapa kelompok berdasarkan satu atau lebih variabel bebas. Melalui metode ini akan dihasilkan peluang dari masing-masing kategori tak bebas yang akan dijadikan sebagai pedoman pengklasifikasian dan suatu pengamatan akan masuk ke dalam variabel tak bebas kategori tertentu berdasarkan nilai peluang yang terbesar.

Ada beberapa metode klasifikasi dengan pendekatan nonparametrik, salah satunya metode klasifikasi berstruktur pohon yang diperkenalkan oleh Leo Breiman, et al (1984). Pada tahun 1984, keempat ilmuwan memperkenalkan metode klasifikasi CART (*Classification And Regression Trees*) yaitu metode pohon regresi dan pohon Klasifikasi. Jika variabel tak bebas yang dimiliki bertipe kategorik maka CART menghasilkan pohon klasifikasi (*classification trees*). Sedangkan jika variabel tak bebas yang dimiliki bertipe kontinu atau numerik maka CART menghasilkan pohon regresi (*regression trees*).

Pada tulisan ini akan dilakukan perbandingan analisis regresi logistik biner dan CART untuk analisis klasifikasi nasabah kredit di suatu bank.

1.2 Permasalahan

Berdasarkan latar belakang tersebut, rumusan masalah dalam penelitian ini adalah bagaimana hasil perbandingan analisis klasifikasi nasabah kredit menggunakan regresi logistik biner dan CART pada kasus data debitur di suatu bank.

1.3 Batasan Masalah

Masalah yang akan dibahas pada tugas akhir ini dibatasi pada variabel-variabel bebasnya. Variabel-variabel bebas yang digunakan dalam penelitian ini adalah jenis kelamin, jenjang pendidikan, jenis pekerjaan, status perkawinan, pendapatan. Variabel tak bebas yang digunakan adalah status kredit nasabah bank. Aturan pemecahan simpul pada CART yang digunakan adalah *twoing criterion*.

1.4 Tujuan Penulisan

Tujuan penelitian dari penulisan tugas akhir ini adalah :

1. Mengetahui hasil analisis regresi logistik biner dan CART (*klasifikasi and reggression trees*) untuk analisis klasifikasi nasabah kredit.
2. Membandingkan akurasi ketepatan klasifikasi menggunakan regresi logistik biner dan CART (*klasifikasi and reggression trees*) pada kasus data debitur di suatu bank.